View Reviews

Paper ID

8

Paper Title

Future Timelines: Extraction and Visualization of Future-related Content From News Articles

Reviewer #1

Questions

1. Overall Rating

Weak Accept

2. Short summary of the paper.

The paper proposes ausystem to extract, sumarrize and present as a temporal timeline content from news collection. To address this, the paper presents an automated system that extracts future-related information about a queried entity from news articles. The approach involves fine-tuning a language model on a diverse dataset, identifying future-related sentences. It uses topic modeling to extract key topics, ranks them by relevance, and employs a temporal tagger to map temporal expressions to sp

3. Three (or more) strong points about the paper.

1) The paper addresses a highly important and relevant topic within the AI community.
2) The system employs state-of-the-art methods for each machine learning task showing its technical rigor.
3) The paper's well-structured and clearly explained methodology greatly contributes to its replicability.

4. Three (or more) improvement points about the paper.

1) Provide a description of practical real-world scenarios in which this problem is relevant. This would help to better understand the practical implications and applications of the research.
2) Enhance the rigor and description of the user evaluation process. A more detailed account of how user feedback and evaluations were conducted.
3) Include a dedicated limitations section. A detailed discussion of potential drawbacks or limitations would provide a more comprehensive view of the research.

5. Presentation

Good: well-structured and mostly clear

6. Originality / Novelty

Limited Novelty

7. Detailed comments. Please number each point and please provide as constructive feedback as possible.

1) It would be beneficial to include real-world scenarios in which the problem addressed in the paper is highly relevant. Providing specific examples or case studies would illustrate how the research findings can be applied in practical settings.

2) A more comprehensive explanation of the user evaluation process is essential. Specify the criteria and metrics used for evaluating user feedback, and detail the methodologies employed for collecting and analyzing this feedback. Clarify the user demographics, such as whether they are news experts or represent a diverse range of users. Specify what aspects they evaluated, whether it was the UX interface, accuracy of the results, or other factors.

3) The paper would gain from the inclusion of a dedicated limitations section. This section should provide insights into scenarios in which the proposed tool may not work effectively. For instance, detailing situations or data types where the tool may face challenges in detecting temporal expressions or outlining potential errors that can occur.

Reviewer #2

Questions

1. Overall Rating

Accept

2. Short summary of the paper.

The authors propose a system to extract future-related phrases related to a query into a news corpus. Their approach involves fine-tuning a language model on a novel constructed corpus of labelled future-related sentences. The user is shown future-related sentences on an interactive timeline.

3. Three (or more) strong points about the paper.

1) The system is a non-trivial end-to-end pipeline hand-designed for the task.

2) The system solves useful problems for users hoping to derive potential future insights about entities.

4. Three (or more) improvement points about the paper.

1) Limited discussion of related work.

5. Presentation

Excellent: careful, logical, elegant, easy to understand

6. Originality / Novelty

Novel

7. Detailed comments. Please number each point and please provide as constructive feedback as possible.

My only concern is, while the authors describe related work, they do not say much about how their system compares, and what other options users might have. They also do not say much about the fundamental limitations of their system (apart from a comment on the speed near the end of section 4), which is an empirical result. It would be good if the authors could identify any design limitations (possibly contrasted with existing work), and how these could be improved.

Reviewer #3

Questions

1. Overall Rating

Accept

2. Short summary of the paper.

The demo paper describe a system that automatically extract
future-related information of a queried entity from news articles. The results are then presented in a nice and intuitive way with the help of an interactive timeline alongside a word cloud depicting the most relevant sentences wrt the query.

3. Three (or more) strong points about the paper.

1. Clever and intuitive data processing pipeline compiled out of state-of-the-art standalone solutions.
2. The data gathering process includes some clever ideas on how to compile a labeled dataset.
3. Easy to understand visualization of the results combining a timeline and a word cloud. The results can be easily navigated and consumed by the user.

4. Three (or more) improvement points about the paper.

1. Related work needs to be more structured with respect to the different topics that you discuss.
2. In section 3.1 you mention that 6,800 sentences are manually labeled but you do not provide more details about it. Who did this labeling? How many entities are involved? How you resolve conflicts? etc. Especially for the case of the News dataset in table 1.
3. In the evaluation section the authors can introduce a small figure depicting the users rating.

5. Presentation

Good: well-structured and mostly clear

6. Originality / Novelty

Novel

7. Detailed comments. Please number each point and please provide as constructive feedback as possible.

1. Overall the demo paper is well written and easy to follow.
2. More clarifications are needed in section 3.1 related to the manual labeling of the dataset especially the News articles.
3. Section 3.2.1 - "Data Retrieval": The authors do not provide any information related to the process of utilizing the user query to fetch related articles. Do you execute a web search and then you crawl the first X results? Do you parallelize the crawling phase/fetching and preprocessing step? Currently this step is not clear to me and additional information needs to be provided.
4. Despite that the UI of the application is easy to use and intuitive it can be improved by considering more suitable libraries such as d3.js
5. I also encourage the authors to include a small paragraph discussing on how to improve the response time of their system. What part of the pipeline can be parallelized and what not. What is the bottleneck and how can be improved?

Reviewer #4

Questions

1. Overall Rating

Weak Accept

2. Short summary of the paper.

This research article presents a framework that automatically extract future related information of a queried entity from news articles. The methodology involves fine-tuning a language model on a novel multisource dataset comprising 6800 annotated sentences to identify future related sentences. Then, topic modeling has been applied to retrieve main topic from the data along with content and subsequently ranked by relevance.

3. Three (or more) strong points about the paper.

1) Innovative approach: introduces a framework that automatically extracts future related information of a queried entity from news articles.
2) Overall Paper is nicely written.
3) Dataset made publicly available for the research community

4. Three (or more) improvement points about the paper.

1) Related work does not describe how the system differs from existing work.
2) The readability of section 3.1 needs to be improved.
3) Add the description of how 6800 sentence has been labeled.

5. Presentation

Good: well-structured and mostly clear

6. Originality / Novelty

Limited Novelty

7. Detailed comments. Please number each point and please provide as constructive feedback as possible.

i) Readability of section 3.1 needs to be improved. For example, coherence is missing between following sentences: “Earlier strategies have frequently relied on extracting data using temporal expressions. However, this approach has its limitations. Predictions can often be intricate, lacking explicit dates or simple temporal cue.”

ii) In Dataset, it is mentioned, “..6,800 manually labeled sentences”, however, how sentences has been labelled (because Table 1 illustrates positive and negative columns) and if annotators were used for labelling then what performance metrics have been used to measure the annotators agreements.

Correction:

I) In abstract, this sentence “We use topic modeling to extract the main topics from the data, which, along with their contents, are subsequently ranked by relevance” needs to be revised.

Typo Error:

Section 1, second paragraph second line, “sift” -> “shift”